BIOTECHNOLOGY
SYSTEMS
BRANCH

RAW SEQUENCE LISTING
ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: *09/005,318B*

Source: *1646*

Date Processed by STIC: *6/23/2000*

RECEIVED

JUL 12 2000

TECH CENTER 1600/2900

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,

2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR FURTHER INFORMATION, PLEASE TELEPHONE MARK SPENCER, 703-308-4212.

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

Checker Version 3.0

The Checker Version 3.0 application is a state-of the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.

Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO).

Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

Checker Version 3.0 can be down loaded from the USPTO website at the following address:
http://www.uspto.gov/web/offices/pac/checker

● ⋃. ЯⱭMℓⱭ ●

1646

```
RAW SEQUENCE LISTING              DATE: 06/23/2000
PATENT APPLICATION:  US/09/005,318B    TIME: 06:20:30

Input Set : A:\401c1.app
Output Set: N:\CRF3\06222000\I005318B.raw
```

**Does Not Comply
Corrected Diskette Needed**

```
                SEQUENCE LISTING
          4 (1) GENERAL INFORMATION:
C--> 6        (i) APPLICANT: Hein, Mich B.
     7                      Hiatt, Andrew C.
     8                      Fitchen, John H.
    10      (ii) TITLE OF INVENTION: NOVEL EPITHELIAL TISSUE TARGETING AGENT
    12     (iii) NUMBER OF SEQUENCES: 140
    14      (iv) CORRESPONDENCE ADDRESS:
    15          (A) ADDRESSEE: SEED IP LAW GROUP PLLC
    16          (B) STREET: 6300 Columbia Center, 701 Fifth Avenue
    17          (C) CITY: Seattle
    18          (D) STATE: Washington
    19          (E) COUNTRY: USA
    20          (F) ZIP: 98104
    22       (v) COMPUTER READABLE FORM:
    23          (A) MEDIUM TYPE: Floppy disk
    24          (B) COMPUTER: IBM PC compatible
    25          (C) OPERATING SYSTEM: PC-DOS/MS-DOS
    26          (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
C--> 28     (vi) CURRENT APPLICATION DATA:
C--> 29          (A) APPLICATION NUMBER: US/09/005,318B
C--> 30          (B) FILING DATE: 09-Jan-1998
    31          (C) CLASSIFICATION:
    33    (viii) ATTORNEY/AGENT INFORMATION:
    34          (A) NAME: Sharkey, Richard G.
    35          (B) REGISTRATION NUMBER: 32,629
    36          (C) REFERENCE/DOCKET NUMBER: 310098.401C1
    38      (ix) TELECOMMUNICATION INFORMATION:
    39          (A) TELEPHONE: (206) 622-4900
    40          (B) TELEFAX: (206) 682-6031


ERRORED SEQUENCES

E--> 472 (2) INFORMATION FOR SEQ ID NO:
   459        (i) SEQUENCE CHARACTERISTICS:
   475           (A) LENGTH: 286 base pairs
   476           (B) TYPE: nucleic acid
   477           (C) STRANDEDNESS: single
   478           (D) TOPOLOGY: linear
   467       (ix) FEATURE:
   468           (A) NAME/KEY: CDS
   469           (B) LOCATION: 1..282
   480       (ix) FEATURE:
   481           (A) NAME/KEY: CDS
   482           (B) LOCATION: 1..282
E--> 488       (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:
```

delete duplicate material (see back page)

```
RAW SEQUENCE LISTING                    DATE: 06/23/2000
PATENT APPLICATION:  US/09/005,318B     TIME: 06:20:30

Input Set : A:\401cl.app
Output Set: N:\CRF3\06222000\I005318B.raw
```

```
        490 GAT CAG AAG TGC AAG TGT GCT CGT ATT ACT TCT AGA ATC ATC CGT AGC        48
        491 Asp Gln Lys Cys Lys Cys Ala Arg Ile Thr Ser Arg Ile Ile Arg Ser
W--> 492 15                  20                  25                  30
        494 TCA GAG GAC CCA AAT GAA GAT ATA GTC GAA CGT AAC ATC CGT ATC ATC        96
        495 Ser Glu Asp Pro Asn Glu Asp Ile Val Glu Arg Asn Ile Arg Ile Ile
W--> 496                     35                  40                  45
        498 GTC CCA CTG AAT AAC CGG GAG AAT ATC TCA GAT CCT ACA AGT CCG TTG       144
        499 Val Pro Leu Asn Asn Arg Glu Asn Ile Ser Asp Pro Thr Ser Pro Leu
W--> 500                 50                  55                  60
        502 CGC ACA CGC TTC GTA TAC CAC CTG TCA GAT CTG TGT AAG AAG TGT GAT       192
        503 Arg Thr Arg Phe Val Tyr His Leu Ser Asp Leu Cys Lys Lys Cys Asp
W--> 504             65                  70                  75
        506 CCA ACA GAG GTA GAG CTG GAC AAT CAG ATA GTC ACT GCG ACT CAA AGC       240
        507 Pro Thr Glu Val Glu Leu Asp Asn Gln Ile Val Thr Ala Thr Gln Ser
W--> 508         80                  85                  90
        510 AAC ATT TGC GAT GAG GAC AGC GCT ACA GAA ACC TGC TAC TGA              282
        511 Asn Ile Cys Asp Glu Asp Ser Ala Thr Glu Thr Cys Tyr
W--> 512 95                  100                 105
        514 ATTC                                                               286
E--> 517 (2) INFORMATION FOR SEQ ID NO: 14:
        519      (i) SEQUENCE CHARACTERISTICS:
        520          (A) LENGTH: 105 base pairs
        521          (B) TYPE: nucleic acid
        522          (C) STRANDEDNESS: single
        523          (D) TOPOLOGY: linear
        525      (ix) FEATURE:
        526          (A) NAME/KEY: CDS
        527          (B) LOCATION: 1..105
        533      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:
        535 GAT CTG TGT AAG AAG TGT GAT CCA ACA GAG GTA GAG CTG GAC AAT CAG        48
        536 Asp Leu Cys Lys Lys Cys Asp Pro Thr Glu Val Glu Leu Asp Asn Gln
W--> 537 95                  100                 105                 110
        539 ATA GTC ACT GCG ACT CAA AGC AAC ATT TGC GAT GAG GAC AGC GCT ACA        96
        540 Ile Val Thr Ala Thr Gln Ser Asn Ile Cys Asp Glu Asp Ser Ala Thr
W--> 541             115                 120                 125
        543 GAA ACC TGC                                                         105
        544 Glu Thr Cys
       1191 (2) INFORMATION FOR SEQ ID NO: 45:
       1193      (i) SEQUENCE CHARACTERISTICS:
       1194          (A) LENGTH: 16 amino acids
       1195          (B) TYPE: amino acid
       1196          (C) STRANDEDNESS:
       1197          (D) TOPOLOGY: linear
       1203      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 45:
E--> 1205
       Ala Ile Gln Asp Pro Arg Leu Phe Ala Glu Glu Lys Ala Val Ala Asp
       2303 (2) INFORMATION FOR SEQ ID NO: 111:
       2305      (i) SEQUENCE CHARACTERISTICS:
       2306          (A) LENGTH: 30 amino acids
       2307          (B) TYPE: amino acid
```

RECEIVED

JUL 12 2000

TECH CENTER 1600/2900

insert hard return so flat number will be under amino acid

RAW SEQUENCE LISTING   DATE: 06/23/2000
PATENT APPLICATION: US/09/005,318B   TIME: 06:20:31

Input Set : A:\401c1.app
Output Set: N:\CRF3\06222000\I005318B.raw

```
         2308          (C) STRANDEDNESS:
         2309          (D) TOPOLOGY: linear
         2314     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 111:
         2316    Val Arg Asp Gln Ala Gln Glu Asn Arg Ala Ser Gly Asp Ala Gly
E--> 2317    1              5           10          15
         2319    Ser Ala Asp Gly Gln Ser Arg Ser Ser Ser Ser Lys Val Leu Phe
E--> 2320    16     20    25        30
         2322 (2) INFORMATION FOR SEQ ID NO: 112:
         2324      (i) SEQUENCE CHARACTERISTICS:
         2325          (A) LENGTH: 25 amino acids
         2326          (B) TYPE: amino acid
         2327          (C) STRANDEDNESS:
         2328          (D) TOPOLOGY: linear
         2334     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 112:
         2336    Val Pro Ser Thr Pro Pro Thr Pro Ser Pro Ser Thr Pro Pro Thr
E--> 2337    1              5                   10          15
         2339    Pro Ser Pro Ser Cys Cys His Pro Arg Leu
E--> 2340    16     20    25
```

(misaligned amino acid nos.)

fix nos

(see item 4 on Error summary sheet)

VERIFICATION SUMMARY   DATE: 06/23/2000
PATENT APPLICATION: US/09/005,318B   TIME: 06:20:32

Input Set : A:\401cl.app
Output Set: N:\CRF3\06222000\I005318B.raw

```
L:6 M:220 C: Keyword misspelled or invalid format, [(i) APPLICANT:]
L:28 M:220 C: Keyword misspelled or invalid format, [(vi) CURRENT APPLICATION DATA:]
L:29 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:30 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:228 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5
L:237 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5
L:433 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:11
L:480 M:201 W: Mandatory field data missing, SeqNo=13, [INFORMATION FOR SEQ ID NO:]
L:480 M:202 E: (16) Value must be an Integer, Data=[]
L:488 M:212 E: (34) Invalid or duplicate Sequence ID Number, Data=[13:]
L:492 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:-1
L:496 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:-1
L:500 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:-1
L:504 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:-1
L:508 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:-1
L:512 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:-1
L:517 M:216 E: (34) Seq.#s missing, 15 thru 13
L:537 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:14
L:541 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:14
L:1148 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:42
L:1152 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:42
L:1205 M:203 E: No. of Seq. differs, LENGTH:Input:16 Found:0 SEQ:45
L:2317 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:111
M:332 Repeated in SeqNo=111
L:2337 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:112
M:332 Repeated in SeqNo=112
L:2486 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:118
L:2489 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:118
L:2640 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:123
L:2646 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:123
L:2649 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:123
```

STATISTICS SUMMARY   DATE: 06/23/2000
PATENT APPLICATION: US/09/005,318B   TIME: 06:20:32

          Input Set : A:\401cl.app
          Output Set: N:\CRF3\06222000\I005318B.raw

Application Serial Number: US/09/005,318B
Alpha or Numeric: Alpha
Application Class: 514
Application File Date: 01-09-1998
Art Unit: 1646
Software Application: PatentIn
Total Number of Sequences: 140
Number of Errors: 8
Number of Warnings: 19
Number of Corrections: 4

MESSAGE SUMMARY

201 W: 1 (Mandatory field data missing)
202 E: 1 ((16) Value must be an Integer)
203 E: 1 (No. of Seq. differs)
212 E: 1 ((34) Invalid or duplicate Sequence ID Number)
216 E: 1 ((34) Seq.#s missing)
220 C: 4 (Keyword misspelled or invalid format)
332 E: 4 ((32) Invalid/Missing Amino Acid Numbering)
341 W: 7 ((46) "n" or "Xaa" used)
336 W: 11 (Invalid Amino Acid Number in Coding Region)